

Reconciling a Diplomatic Edition of the Vidin Miscellany with Critical Editions of its Texts A New Approach to Normalisation

Approaches To The Editing of texts with a Multilingual Tradition (ATTEMPT)
19-20 December 2013, King's College London

David J. Birnbaum and Lara Sels

In the Slavistic scholarly tradition, the boundaries between diplomatic editions of manuscripts and critical editions of texts have often been blurred, not only because of particulars in the historical development of the field, but also because of specific methodological problems that inhere in the textual traditions themselves. The case of the *Vidin Miscellany* (henceforth *VM*) will be used to illustrate some of those problems and possible solutions. As a *codex unicus* of an interesting, fourteenth-century compilation of translated female Saints' *Vitae*, the miscellany presents its editors with many challenges. We will argue that a complete edition of *VM* should incorporate at least three perspectives, viz. that of *VM* 1) as a document (diplomatic edition of the manuscript), 2) as a collection of texts (critical editions of the individual entries), and 3) as an entity made up of texts that travel together in transmission, viz. a text in its own right (critical edition of the Vidin collection, based on the *VM* as a unique text witness). These three perspective can inform one another only if full weight is given to critical research into the origins and the transmission of both the individual entries (including their Greek source texts) and the collection as a whole.

As we are dealing with manuscript texts in their material particularity on the one hand, and with these texts as abstractions extrapolated from their various attestations on the other, we need to work with both diplomatic transcriptions and normalised text files. In view of the presence of various linguistic recensions within one textual tradition and in view of the high degree of orthographic variation in the manuscripts, normalisation is a major and problematic issue in palaeoslavistic editorial practice, not only for the establishment and final presentation of the text, but also in the preparatory stages of the edition. The normalization process is particularly important where the textual variants are to be collated with the help of computational alignment tools because computers are particularly effective at exact string matching, and may not easily distinguish significant from superficial variation in surface forms. Our current research seeks to use diplomatic transcriptions directly as the input for the collation of variants, so that the researcher does not have to prepare separate transcriptions for the two purposes or edit one manually to turn it into the other. Because the available open-source applications with which we are acquainted (e.g., Juxta, CollateX) are not easily able to collate texts with the degree of orthographic variation that we find in our manuscripts, we are developing, in Python, our own open-source tool, using a modification of the Soundex algorithm for pattern matching and an implementation of suffix tries for the efficient identification and alignment of corresponding variants in our witnesses.